

Guaranteeing Service Availability in SLAs; a Study of the Risk Associated with Contract Period and Failure Process

Andrés J. González and Bjarne E. Helvik, *Member, IEEE*.

Abstract— Service Level Agreements (SLAs) are a common means to define the obligations of network/service providers and users in business relationships. The terms that define the guaranteed availability for a given period are an important element of these contracts. The appropriate selection values is difficult due to the large number of variables involved, the complexities of the network and service provision and the computational challenge posed by the transient solution, as opposed to a steady state, that is needed. A common policy taken to solve it, is using the steady state availability as a reference. Nevertheless this simplification may put on risk the contract fulfillment as stochastic variation of the measured availability is significant over a typical contract period. This paper analyzes the relevance that the interval availability analysis has on SLAs, and provides suggestions to the network providers on the selection of adequate availability guarantees. The interval availability of unprotected and shared protected connections is studied under exponential and Weibull failure and repair distributions. It is observed that for a single path scenario, a small reduction of the guaranteed availability below the steady state value improve the probability to meet the requirements considerably. The same is the case for connections with shared backup protection. However performing this analysis in the transient domain is quite demanding. Hence, to simplify it, it is proposed to obtain the steady state results and introduce a *safeguard factor* to control that the availability guarantee is met. For the Weibull distributed times between failures, where the shape factor is less than one, as observed in operational networks, the probability of meeting a guaranteed availability over a finite contract period, decrease more radically than for the commonly assumed Poisson failure process. This increases the importance of making a transient analysis.

Keywords — Network dependability, failure characterization, Weibull distribution, SLA definition, risk in SLAs.

I. INTRODUCCIÓN

LAS EMPRESAS de Telecomunicaciones y sus respectivos clientes utilizan los Acuerdos de Nivel de Servicio (SLA) para definir niveles específicos de calidad de servicio QoS, donde disponibilidad es un elemento significativo.

Centre for Quantifiable Quality of Service in Communication Systems, Centre of Excellence" appointed by The Research Council of Norway, funded by the Research Council, NTNU and UNINETT. <http://www.q2s.ntnu.no>.

A.J. Gonzalez es investigador del departamento de Telemática y del Centro Q2S-NTNU, O.S Bragstads plass 2E, 7491 Trondheim, Norway (correo e.: andresgm@q2s.ntnu.no)

B.H. Helvik es vicedecano de investigación de la Facultad de Información y Tecnología NTNU, O.S Bragstads plass 2E, 7491 Trondheim, Norway (correo e.: bjarne@q2s.ntnu.no)

El incumplimiento de alguno de los valores pactados genera grandes consecuencias en la reputación, la economía y el desempeño de las dos partes implicadas.

Esta investigación estudia como definir los parámetros de disponibilidad en SLAs teniendo en cuenta el comportamiento transitorio de esta durante un periodo fijo en el cual el contrato este vigente. Suponiendo que las propiedades de Markov se cumplen en los procesos de fallas que afectan los componentes de red, la *disponibilidad transitoria* (disponibilidad durante un periodo de tiempo) puede ser obtenida mediante métodos numéricos usando técnicas de uniformización, tal como es presentado en [15], [17], [8]. Es de especial interés analizar la probabilidad que tiene un proveedor de servicios de red de alcanzar la disponibilidad estipulada en el contrato α , la cual será definida como *probabilidad de éxito del SLA*, en una red cuyas conexiones se ven afectadas por fallas y reparaciones respectivamente.

Este problema fue inicialmente tratado para sistemas generales en [10], donde fue observado que si α es mayor que la disponibilidad en estado estable (A) que tiene un servicio, la probabilidad de éxito del SLA se reduce progresivamente hasta llegar a 0, lo cual puede ser intuitivo. Sin embargo también fue observado que existe un riesgo considerable, incluso cuando $\alpha < A$. En [16] se sugiere de forma general como usar el concepto de disponibilidad transitoria para hacer una planeación económica en donde se maximice las ganancias en un contexto de SLAs. Un artículo más reciente [3] estudió el comportamiento transitorio bajo el uso de estrategias adaptativas de administración de recursos para controlar las posibles multas y la coherencia en los compromisos adquiridos.

Un tema importante en este contexto, son las conexiones que utilizan backups compartidos. La popularidad de esta técnica ha crecido en los últimos años dado que ofrece un considerable incremento en la disponibilidad y mantiene a su vez un eficiente uso de los recursos de red [6], [5], [2]. En relación a este tema, la mayoría de las investigaciones han modelado la disponibilidad a través de cotas para aproximar la solución asintótica. Esto es necesario debido a la complejidad que implica el encontrar una solución exacta dadas las dependencias que existen entre los diferentes procesos cuando los links son compartidos [12], [9], [11], [13], [4]. Este artículo extiende esos resultados mediante el análisis del comportamiento transitorio de la disponibilidad en dichos

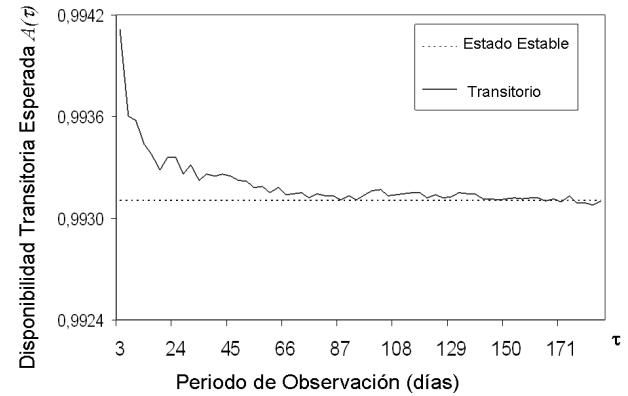
escenarios. Un descubrimiento inicial de nuestro estudio ilustra que es más importante encontrar una solución aproximada asintótica que una más precisa, ya que esta segunda requiere un enorme esfuerzo computacional con un error no controlado dado el riesgo que representa tener valores de α cerca de A .

Durante los últimos años ha sido mostrado que los tiempos entre fallas o entre reparaciones de un link pueden ser modelados más precisamente usando la distribución Weibull en vez del uso de distribuciones exponenciales, las cuales de echo son asumidas más comúnmente en el plano teórico. Esto se debe a que el comportamiento real de los dispositivos de red posee un número más elevado de ocurrencias de corto y de largo plazo, e.g. [14]. Por tanto la probabilidad de éxito será analizada bajo procesos de fallas y reparaciones que siguen una distribución Weibull. Fue detectada una relación interesante entre el parámetro de forma (β) de una función Weibull y la probabilidad de éxito, lo cual resalta la importancia del uso de modelos más realistas con el fin de definir términos más responsables de disponibilidad en un SLA.

Herramientas para el análisis numérico de la disponibilidad transitoria son encontradas solo para el caso de sistemas triviales (un solo elemento) o para sistemas generales que cumplen las propiedades de Markov. Dado que nuestras posiciones son más realistas acá usamos simulación para obtener resultados haciendo uso de la plataforma Möbius [1], [18]. En todos los escenarios usados las simulaciones fueron ejecutadas usando intervalos de confianza de 0.1 para tener un nivel de confianza del 99%.

Con el fin de que los resultados obtenidos en dicha simulación sean fácilmente aplicables, se introdujo el concepto de *Factor de Corrección* σ que permite el uso de valores de disponibilidad asintóticos (los cuales son más fáciles de obtener), pero a la vez teniendo en cuenta el efecto estocástico que tiene el comportamiento transitorio de la disponibilidad, dado que los SLA son definidos solo para un periodo finito de tiempo. De este modo la disponibilidad transitoria ofrecida en el contrato podrá ser cumplida con una probabilidad alta.

Este documento está organizado de la siguiente manera. La Sección II introduce y explica los conceptos relacionados con la distribución probabilística de la disponibilidad transitoria y su relación con SLAs. En la Sección III se estudia la probabilidad de éxito de una conexión con backup compartido, i.e. la probabilidad que la disponibilidad observada sea acorde a las promesas realizadas. La Sección IV presenta el efecto de tener procesos de fallas y reparaciones que sigan una distribución Weibull. La Sección V discute como pueden ser obtenidas promesas de disponibilidad que sean seguras. Finalmente la Sección VI concluye este artículo.



(a) Disponibilidad Transitoria Esperada

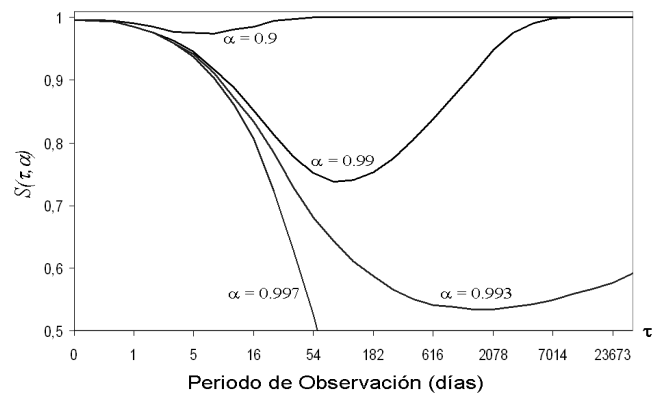
(b) Probabilidad de éxito bajo diferentes valores de α para un sistema con $A=0.99311514$

Figure 1: Comportamiento transitorio de la disponibilidad

II. PROBABILIDAD DE EXITO EN SLAS

Una *conexión de red* es definida como un grupo de links interconectados con el fin de proveer un servicio de un punto inicial a uno final. Mantener dicha conexión en estado operacional (funcionando) es de vital importancia para ofrecer una buena calidad de servicio. Su desempeño como una función dependiente del tiempo puede ser modelado utilizando el proceso estocástico $\hat{I}(t)$ definido como se enuncia a continuación:

$$\hat{I}(t) = \begin{cases} 1 & \text{Conexión esta activa en } t \\ 0 & \text{En otro caso} \end{cases} \quad (1)$$

Específicamente en un escenario donde se utilizan SLAs, es importante estudiar la disponibilidad transitoria $\hat{A}(\tau)$, la cual es una variable estocástica que mide el tiempo total que una conexión ha estado en funcionamiento durante un periodo τ y que puede ser descrita por la siguiente ecuación:

$$\hat{A}(\tau) = \frac{1}{\tau} \int_0^{\tau} \hat{I}(t) dt \quad (2)$$

Cuando se estudia el comportamiento transitorio de una conexión, es usualmente evaluada la disponibilidad transitoria esperada (promedio), definida como $A(\tau) = E[\hat{A}(\tau)]$.

Por ejemplo la Fig. 1(a) muestra que una conexión que tiene cinco links que fallan de forma independiente acorde a una distribución exponencial negativa con media de un año y tiempos de reparación promedio de 12 horas, $A(\tau)$ converge en pocas semanas hacia su estado estable A definido de la siguiente forma.

$$A = \lim_{\tau \rightarrow \infty} A(\tau) \quad (3)$$

Sin embargo el concepto de $A(\tau)$ no es suficientemente descriptivo debido a que no considera los diferentes momentos estadísticos que pueden tener una influencia crítica en la calidad de la conexión. Dado un contexto en el cual un SLA garantiza una disponibilidad α , y teniendo en cuenta todos los momentos de la distribución estocástica de la disponibilidad transitoria, es de gran importancia evaluar que probabilidad existe de que la disponibilidad ofrecida sea igual o mayor a la prometida en un periodo de observación τ . Dicho valor será definido como la *probabilidad de éxito*:

$$S(\tau, \alpha) = P[\hat{A}(\tau) \geq \alpha] \quad (4)$$

Adicionalmente el *riesgo* de un SLA será definido como la probabilidad de que un valor de disponibilidad prometido no sea cumplido, lo cual puede ser expresado como $1 - S(\tau, \alpha)$.

Existen algunos métodos numéricos que mediante el uso de técnicas de uniformización pueden calcular $S(\tau, \alpha)$ para sistemas Markovianos [15], [17], [8]. Adicionalmente, para el caso de un componente individual que tiene procesos de error y mantenimiento con distribuciones generales, pueden ser obtenidos resultados puntuales mediante el uso de teorías de procesos renovables (renewal theory) [7]. Sin embargo, debido a que en este estudio se modelan conexiones con varios componentes que muchas veces no cumplen las propiedades de Markov i.e. procesos de fallas y reparación con distribución Weibull, los valores de $S(\tau, \alpha)$ serán calculados mediante el uso de simulación.

La probabilidad de éxito en función del periodo de observación τ fue tratado por primera vez por Goyal et. al. [10]. $S(\tau, \alpha)$ básicamente tiene dos comportamientos que dependen de los valores de A y α (Ver Fig. 1(b)). En el primer caso cuando $\alpha \leq A$ la curva desciende durante algunos periodos pero a medida que dicho periodo aumenta, esta converge hacia un valor de éxito igual a 1. La duración y valor del riesgo máximo alcanzado, depende de la proporción que exista entre A y α . En el segundo caso cuando $\alpha > A$ la probabilidad de éxito decrece continuamente hasta alcanzar un valor de 0.

Teniendo en cuenta estos factores, en este artículo se define una promesa como *responsable* si durante un periodo τ la disponibilidad prometida α es menor que la disponibilidad asintótica A .

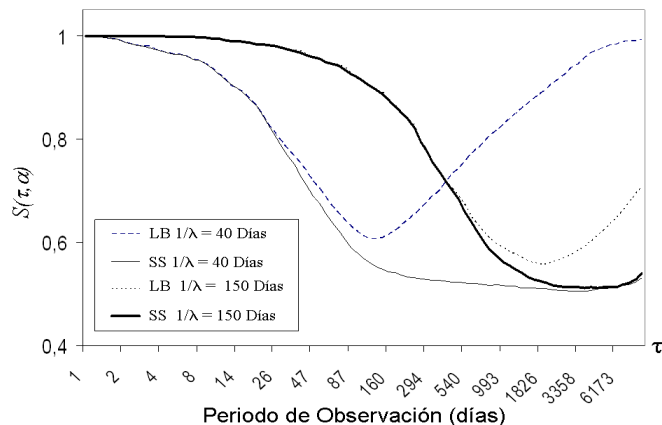


Figure 2: Probabilidad de éxito con $\alpha =$ Cota Inferior (LB) y $\alpha \approx$ Cerca al valor de disponibilidad en estado estable (SS) bajo links de diversas calidades

El ejemplo mostrado en la Fig. 1(b) fue elaborado utilizando un escenario de cinco links independientes que tienen procesos de fallas distribuidos Poisson con un valor esperado de un año y tiempos de reparación i.i.d con distribuciones exponenciales negativas y valor esperado de 12 horas. La conexión punto a punto tiene una disponibilidad en estado estable de $0,99315$. Se puede observar también que el riesgo puede llegar a ser grande y que esto solo puede ser evitado si se escogen tiempos de observación τ muy cortos o muy largos.

Sin embargo, estos periodos en los cuales hay cierta seguridad en el cumplimiento del contrato están fuera de rangos reales en los cuales un SLA es estipulado (e.g 1 día o 20 años). Por consiguiente es relevante buscar otro tipo de soluciones como por ejemplo el uso de garantías α menores al valor de la disponibilidad asintótica del sistema A . Es importante destacar que cuando $A = \alpha$, $S(\tau, \alpha)$ cae rápidamente alcanzando un riesgo de 0.5 para cualquier periodo τ .

III. CONEXIONES CON PROTECCION COMPARTIDA

Las empresas de redes de telecomunicaciones han venido implementando mecanismos de protección para cumplir y garantizar la disponibilidad prometida mediante la reservación de recursos adicionales que puedan ser usados cuando los principales fallen. Esto incluye la asignación de canales principales (W) los cuales son dedicados y canales compartidos usados como backup (B). Nuestro estudio en esta sección asume que los links que hacen parte del backup pueden ser compartidos por varias conexiones. Este mecanismo es conocido como *protección compartida* y permite combinar conceptos tan importantes como protección y uso eficiente de los recursos de red.

Bajo este enfoque, ambos canales (principal y backup) no deben tener ningún recurso en común. Cualquier conexión utiliza por defecto el canal principal, pero si este es afectado por algún error, el backup debe ser usado en remplazo (siempre y cuando este también este funcionando). Adicionalmente, debido a que este remplazo no es dedicado la disponibilidad de dicha conexión también va a depender del

uso potencial que otras conexiones hagan de dichos recursos compartidos.

Se podría decir entonces que la disponibilidad total de una conexión depende de sus propios recursos, pero a su vez de la disponibilidad del canal principal de otras conexiones con las cuales se comparte el backup. Por consiguiente el cálculo de la disponibilidad en estos escenarios es mucho más difícil que para el caso de conexiones sin protección o con protección dedicada. La complejidad es aun mayor si se tiene en cuenta que las conexiones en una red son continuamente establecidas o desinstaladas, haciendo dinámico el número de conexiones con las cuales se comparten recursos [4]. Por esta razón nosotros utilizaremos cotas con el fin de obtener valores asintóticos aproximados de una manera más eficiente.

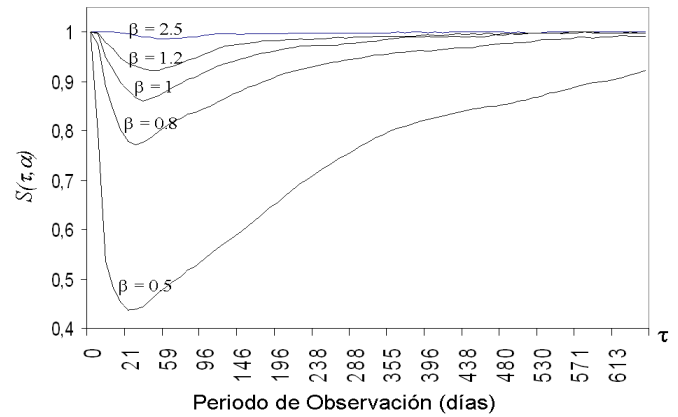
Para calcular la disponibilidad asintótica A^{C_n} de una conexión C_n que usa protección compartida, es necesario tener en cuenta: La indisponibilidad del canal principal el cual contiene un número dado de links con respectivas disponibilidades individuales asintóticas A_i , segundo la indisponibilidad del canal de backup que contiene un numero de links l_j , cada uno con disponibilidad asintótica A_j . Finalmente se considera la disponibilidad de los canales principales (W) del grupo de conexiones C_s que comparten recursos con C_n . Esto se puede definir usando la siguiente ecuación:

$$A^{C_n} \geq 1 - \left(1 - \prod_{\forall i | L_i \in W} A_i \right) \cdot \left(1 - \prod_{\forall j | L_j \in B} A_j \cdot \left[\prod_{\forall j | C_s} A^{C_s} \right] \right) \quad (5)$$

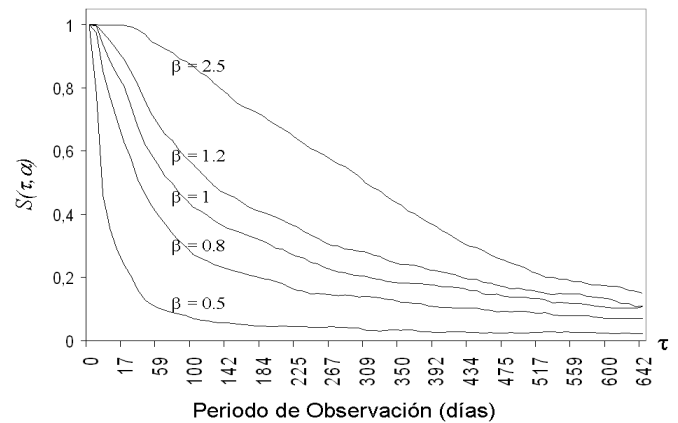
El análisis es mucho más complejo cuando se trata de obtener valores no asintóticos (transientes). Por consiguiente, para este caso es más eficiente hacer uso de simulación con el fin de obtener valores específicos. El escenario utilizado para generar la Fig. 2, contiene dos conexiones, en donde cada una tiene un canal principal y un backup disyunto. Cada uno de los respectivos backups comparte un link que es usado en caso de que algún canal principal falle. Dicho link compartido puede ser usado por la conexión que lo necesite primero, si este esta como tal funcionando y no ha sido solicitado previamente, de otro modo dicha conexión tendrá que esperar a que se reestablezca el funcionamiento del canal principal de la conexión que esta haciendo uso de el.

La Fig. 2 muestra los resultados obtenidos en la simulación, para links con dos tiempos promedio diferentes entre fallas de los links. El primero usando una garantía α inferior pero lo mas cercana posible a la disponibilidad en estado estable A del sistema ($\alpha \approx A$) y segundo usando un valor de α igual a la cota inferior obtenida con la ecuación 5.

En general se observó que las propiedades de α se mantienen para los escenarios donde existen conexiones con backups compartidos. Tal como se esperaba, de acuerdo con los resultados obtenidos en la sección previa, la cota inferior es siempre una decisión más segura para un proveedor de red en el contexto de un SLA.



(a) Probabilidad de éxito bajo diferentes valores de β en procesos de fallas en links (β), cuando $A \geq \alpha$



(b) Probabilidad de éxito bajo diferentes valores de β en procesos de fallas en links (β), cuando $A < \alpha$

Figure 3: $S(\tau, \alpha)$ Behavior under Weibull Process

Especialmente esta ventaja es mayor cuando los links no son de buena calidad (alto promedio de fallas). A pesar de que escoger la cota inferior es mas seguro, la probabilidad de éxito y el riesgo de fallar el contrato puede llegar a ser significativa y por tanto aun es necesario mejorar dicho valor.

IV. ANÁLISIS WEIBULL

En medidas hechas en sistemas operacionales se ha observado la existencia de procesos de fallas y reparaciones con alta concentración de tiempos entre-fallas o muy cortos o muy largos e.g. [14], [19], comportándose de manera mas abrupta que la descrita por un proceso con distribución exponencial negativa. Por consiguiente estos procesos pueden ser modelados de una manera mas precisa haciendo uso de distribuciones Weibull con parámetro β (forma) menor a 1 i.e. nosotros asumimos tiempos entre-eventos T_x que se distribuyen de la siguiente forma:

$$P(T_x > t) = 1 - F_{Tx}(t) = e^{-\left(\frac{t}{\theta}\right)}, \forall t \geq 0 \quad (6)$$

donde θ es el parámetro de escala de la función Weibull.

Es importante resaltar que cuando β es igual a 1, la distribución se convierte en exponencial, siendo los resultados obtenidos en la sección anterior un caso particular de la diferente gama de distribuciones Weibull.

Como se explicó anteriormente, la probabilidad de éxito es obtenida para conexiones sin protección y con protección compartida. Los valores θ son configurados de modo que los tiempos promedio entre fallas sea el mismo (un año) para todos los diferentes valores de β usados. Los tiempos de reparación usados también se distribuyen Weibull, con parámetro de forma menor a uno y un valor esperado de 12 horas para todos los casos. Por consiguiente los valores asintóticos son los mismos para todos los casos mostrados en esta sección.

La Fig. 3 muestra algunos resultados en donde el valor de α es igual a 0.98 ilustrando un ejemplo de un contrato *responsable* (Fig. 3(a)) y α igual a 0.997 para un ejemplo en el cual α es mayor que A (Fig. 3(b)). Como puede ser observado, para distribuciones exponenciales el riesgo no es tan crítico como cuando el parámetro de forma de la distribución de fallas (β_f) es inferior a uno. La Fig. 3(a) muestra que si los posibles valores de β_f son muy pequeños, la probabilidad de éxito $S(\tau, \alpha)$ puede reducirse de una manera pronunciada y amplia. Adicionalmente, los tiempos necesarios para retornar a valores aceptables de riesgo son mucho más largos que para el caso de exponencial simple ($\beta_f = 1$), y pueden exceder los tiempos convencionales de duración de un SLA. Cuando $A < \alpha$, se puede observar en la Fig. 3(b) que el efecto de β_f es también significativo y puede potencializar y hacer más rápido el fracaso del SLA.

El efecto del parámetro de forma Weibull en los procesos de reparación (β_r), también fue evaluado. Se encontró que si dicho valor es menor que 1, se suaviza el efecto causado por el efecto de β_f en el proceso de fallas y por tanto el riesgo se reduce un poco. Sin embargo este efecto no es significativo bajo condiciones convencionales de redes de comunicaciones (tiempos de fallas del orden de meses y reparaciones del orden de horas) debido a que las fallas se convierten en procesos dominantes sobre el resultado de la disponibilidad. Los resultados descritos en esta sección muestran la importancia de una caracterización apropiada de los procesos de fallas, ya que los impactos que esta tiene en el éxito de un SLA pueden ser demasiado relevantes.

V. CUMPLIENDO GARANTIAS DE DISPONIBILIDAD

El análisis realizado en la sección anterior mostró que el comportamiento estocástico de la disponibilidad transitoria puede poner en riesgo el cumplimiento de los contratos de disponibilidad. Por esta razón deben ser tomadas ciertas medidas que contrarresten este peligroso efecto.

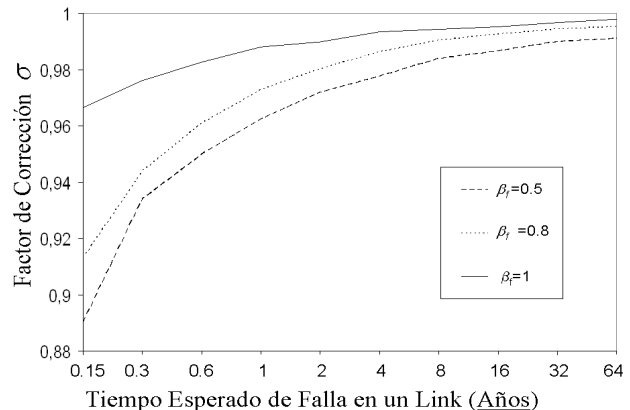


Figure 4: Factor de Corrección σ necesario para obtener un 99% de Probabilidad de Éxito

Partiendo del hecho de que para un proveedor de red es extremadamente difícil tener una probabilidad de éxito del 100%, deben ser planeados valores de riesgo pequeños teniendo en cuenta los costos comerciales y técnicos que esto implica.

Con este fin, será escogido un valor de disponibilidad a garantizar en un SLA α^* que permita tener un riesgo controlado.

Para obtener dicho valor será definido un factor de seguridad σ el cual permitirá hacer el ajuste a la disponibilidad del sistema en estado estable A con el fin de obtener el nivel de éxito deseado el cual puede ser definido basado en las consecuencias que implican la violación de los parámetros contratados, i.e.

$$\alpha^* = \sigma \cdot A \quad (7)$$

De esta forma es posible tener una solución con la facilidad y agilidad que implica la solución asintótica, pero a su vez con el realismo y la seguridad que ofrece tener en cuenta el comportamiento transitorio.

Con el fin de tener una idea más clara de los valores implicados para tener una adecuada definición del SLA, se realizaron varias simulaciones en donde se evalúa la disponibilidad transitoria de una conexión con cinco links que fallan y son reparados de forma independiente siguiendo una distribución Weibull.

En la Fig. 4 pueden ser observados los resultados que muestran la tendencia del factor de seguridad en un SLA de un año de duración. Los valores de beta mostrados allí, son correspondientes a los procesos de fallas (β_f).

También es importante tener en cuenta el efecto que tiene la calidad de cada link de manera individual sobre σ . En la Fig. 4 se puede apreciar que cuando se quieren obtener valores de corrección cercanos a 1 (menos desviación) se necesitan links con una frecuencia de fallas demasiado baja (links de alta calidad). En la práctica, tener links con estas características puede implicar un esfuerzo gigante y por

conseguido en las redes actuales es más realista prever factores de corrección σ un poco más flexibles.

Basado en la Fig. 4, un caso particular que puede ilustrar mejor el efecto que tiene el uso del factor de corrección en un SLA es el siguiente: Asumiendo un contrato de un año de duración (365 días) sobre una conexión para la cual se calcula una disponibilidad en estado estable (A) de 363 días, si los links que hacen parte de la conexión tienen procesos de fallas que siguen una distribución Weibull con promedio de un año y β_f igual a 0.5. Asumiendo que se desea tener una probabilidad de éxito de 99%, habría que estipular en el SLA con 16 días de indisponibilidad (14 días más que lo esperado con A).

Si esta diferencia se analiza sobre un modelo de negocios, se deben considerar las diferentes variables que intervienen en la imagen comercial del proveedor. Por un lado prometer altos niveles de disponibilidad tiene un impacto bastante benéfico a corto plazo, pero por el otro lado si el proveedor promete una disponibilidad que seguramente no va a poder cumplir, las consecuencias a su reputación pueden ser muy fuertes, especialmente a largo plazo.

Como fue mostrado anteriormente, hay muchas variables que deben ser tenidas en cuenta cuando se definen los términos de disponibilidad. Desde un punto de vista comercial se destacan tres variables. Primero el costo que implica la instalación, configuración y asignación de recursos de red con el fin de proveer conexiones de alta calidad, que incluyen la selección de infraestructura altamente confiable y la asignación de protecciones adecuadas. Segundo el impacto comercial que puede tener ofrecer niveles de disponibilidad no muy atractivos y que pueden ser mejorados fácilmente por proveedores de red de la competencia (como se mostró anteriormente, desde el punto de vista matemático esta diferencia puede ser del orden de días). Finalmente la tercera variable es el costo que genera los reembolsos causados por el incumplimiento del contrato, junto con los impactos negativos que esto puede tener a largo plazo en la imagen del proveedor.

Una de las ideas detrás del factor de corrección y los otros análisis realizados en este estudio es ofrecer herramientas que le permitan al proveedor balancear de una manera más clara los tres factores anteriormente mencionados con el fin de ofrecer calidad de servicio y obtener la mayor ganancia posible. La configuración exacta de las variables puede variar de un proveedor a otro basado en las características particulares que las redes de estos pueden tener.

VI. CONCLUSIONES

Este artículo demuestra la importancia de tener en cuenta la distribución estocástica de la disponibilidad transitoria para la definición y especificación de un SLA.

Fue mostrado que cuando un proveedor ofrece una disponibilidad igual a la dada por el estado estable, existe una caída hasta 0.5 en $S(\tau, \alpha)$ que se extiende por un tiempo infinito. Con el fin de tener una garantía en el SLA que pueda ser cumplida, esta debe ser especificada por debajo del valor de la disponibilidad en estado estable. Para el caso de

protección compartida se confirmó que la cota inferior obtenida en estudios anteriores es útil y suficientemente para ser usada en la definición de una garantía apropiada alfa. Se demostró que $S(\tau, \alpha)$ tiene un comportamiento similar bajo procesos de fallas y reparación que tienen una distribución Weibull, pero adicionalmente se encontró que el parámetro de forma β tiene un efecto considerable en el éxito del SLA. Para valores de β_f menores a 1 (valor usual en redes actuales) se encontró que el riesgo aumenta considerablemente y puede extenderse por un periodo que puede superar los tiempos típicos de duración de un SLA.

Finalmente se demostró que es posible controlar la probabilidad de éxito usando un procedimiento sencillo y ágil.

REFERENCIAS

- [1] Mobius Manual. Version 2.2.1. Nov. 2008.
- [2] A. Mykkeltveit and B.E. HelvikE, "Comparison of Schemes for Provision of Differentiated Availability-guaranteed Services Using Dedicated Protection," in *Proc. Seventh International Conference on Networking (ICN)*, Apr. 2008.
- [3] A. Mykkeltveit and B.E. Helvik, "Adaptive management of connections to meet availability guarantees in SLAs," in *Proc. IM 2009 Mini-Conference*, Jun. 2009.
- [4] A. Mykkeltveit and B.E. HelvikE, "On Provision of Availability Guarantees Using Shared Protection," in *Proc. 12th Conference on Optical Network Design and Modelling (ONDM)*, Mar. 2008.
- [5] A. Mykkeltveit and B.E. HelvikE, "Provision of connection-specific availability guarantees in communication Networks," in *Proc. 6th International Workshop on Design of Reliable Communication Networks (DRCN)*, Oct. 2007.
- [6] P. Cholda, A. Mykkeltveit, B.E. Helvik, O.J. Wittner, and A. Jajszczyk, "A survey of resilience differentiation frameworks in communication Networks," *IEEE Communications Surveys & Tutorials*, vol. 9, pp. 32-55, Fourth Quarter 2007.
- [7] D.R. Cox, *Renewal Theory*, Methuen 1967.
- [8] E. De Souza E Silva and H.R. Gail, "Calculating Cumulative Operational Time Distributions of Repairable Computer Systems," *IEEE Trans. on Computers*, vol. C-35, iss. 4, pp. 322-332, Apr. 1986.
- [9] A. Fumagalli, M. Tacca, F. Ungbhvary, and A. Farago, "Shared path protection with differentiated reliability," in *Proc. IEEE International Conference on Communications ICC*, 2002.
- [10] A. Goyal, and A.N. Tantawi, "A measure of guaranteed availability and its numerical evaluation," *IEEE Trans. on Computers*, vol. 37, pp. 25-32, Apr. 1988.
- [11] H. Pin-Han, and J. Tapolcai, and A. Haque, "Spare Capacity Reprovisioning for Shared Backup Path Protection in Dynamic Generalized Multi-Protocol Label Switched Networks," *IEEE Trans. On Reliability*, vol. 57, pp. 551-563, Apr. 2008.
- [12] L. Hongbin, L. Lemin, and Y. Hongfang, "Routing Connections With Differentiated Reliability Requirements in WDM Mesh Networks," *IEEE/ACM Trans. On Networking*, vol. 17, no 1, pp. 253-266, Feb. 2009.
- [13] H. Changcheng, L. Minzhe, and A. Srinivasan, "A Scalable Path Protection Mechanism for Guaranteed Network Reliability Under Multiple Failures," *IEEE Trans. on Computers*, vol. 56, no. 2, pp. 254-267, Jun. 2007.
- [14] A. Markopoulou, G. Iannaccone, S. Bhattacharyya, C. Chuah, Y. Ganjali, and C. Diot, "Characterization of Failures in an Operational IP Backbone Network," *IEEE/ACM Trans. On Networking*, vol. 16, pp. 749-762, Aug. 2008.
- [15] D.A.A. Mello, A. Dominic, A. Schupke, and H. Waldman, "A matrix-based analytical approach to connection unavailability estimation in shared backup path protection," *IEEE Communications Letters*, vol. 9, no.9, pp. 844-846, 2008.
- [16] D.A.A. Mello, G.S. Quiterio, A. Dominic, A. Schupke, and H. Waldman, "Specification of SLA survivability requirements for optical path protected connections," *Optical Fiber Communication Conference*,

2006, and, *The 2006 National Fiber Optic Engineers Conference. OFC 2006*.

- [17] G. Rubino, and B.A. Sericola, "Interval availability analysis using denumerable Markov processes: application to multiprocessor subject to breakdowns and repair," *IEEE Trans. on Computers*, vol. 44, iss. 2, pp. 286-291, Feb. 1995.
- [18] W. H. Sanders, and W. D. Obal, "Dependability evaluation using UltraSAN," in *Proc. Twenty-Third International Symposium on Fault-Tolerant Computing FTCS-23. Digest of Papers*, pp. 286-291, Jun. 1993.
- [19] The Norwegian Research Network UNINETT. Downtime statistics. Available at: <http://drift.uninett.no/downs/>.



Andrés J. González nació en Tunja, Colombia, el 25 de Enero de 1981. Se graduó en la Universidad Javeriana como Ingeniero Electrónico y de la Universidad Nacional de Colombia como Master en Telecomunicaciones. Actualmente es estudiante de Doctorado e investigador de la Universidad Noruega de Ciencia y Tecnología (NTNU) en el Centro de Calidad de Servicio Cuantificable de Sistemas de Comunicación (Q2S). Sus investigaciones incluyen evaluación de disponibilidad, sistemas tolerantes a fallas, así como

temas relacionados a arquitecturas de redes de telecomunicaciones.



Bjarne E. Helvik (M'85) se graduó como MSc en Tecnología del Instituto Noruego de Tecnología (NTH), Trondheim, Noruega en 1975, y recibió el grado de Dr. Techn. del NTH en 1982.

Actualmente es profesor en la Universidad Noruega de Ciencia y Tecnología (NTNU) en el Departamento de Telemática, y es uno de los fundadores de uno de los principales Centros de Excelencia de Noruega (CoE), el Centro de Calidad de Servicio Cuantificable de

Sistemas de Comunicación (Q2S). Desde 2009 ha sido Vicedecano de investigación de la Facultad de Información y Tecnología de la NTNU. Sus investigaciones incluyen QoS, evaluación de disponibilidad, sistemas tolerantes a fallas, así como temas relacionados a arquitecturas de redes de telecomunicaciones.

El profesor Helvik es activo en varias cooperaciones internacionales y varios comités de programas científicos y técnicos.